12-2018

# A Fine Scale Population Genetics Study of the Rare *Helianthus verticillatus*

Tyler Patton Edwards
*University of Tennessee*, tedwar25@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Tyler Patton Edwards entitled "A Fine Scale Population Genetics Study of the Rare *Helianthus verticillatus*." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Entomology and Plant Pathology.

Denita Hadziabdic-Guerry, Major Professor

We have read this thesis and recommend its acceptance:

Bonnie Ownley, Phillip Wadl, Alan Windham

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

# A Fine Scale Population Genetics Study of the Rare *Helianthus verticillatus*

A Thesis Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Tyler Patton Edwards
December 2018

## *Dedication*

I would like to dedicate this to my grandparents who forced me to garden as a child and taught me about plants against my will. I may have not realized at the time, but you showed me something I truly love and believe can make the world a better place.

# *Acknowledgements*

I would like to thank all of my family and friends who helped keep me sane during graduate school. I would like to thank Malcolm Hodges of The Nature Conservancy and Kyla Cheynet of Weyerhaeuser Corporation for allowing me and helping me to conduct this research on their land, as well as the USDA for funding this project. I would also like to give a special thanks to Dr. Denita Hadziabdic-Guerry, my committee and all of the Trigiano lab for putting up with me and believing in me even when I didn't. Without all of you this would not have been a possibility.

# *Abstract*

*Helianthus verticillatus*, the whorled sunflower, is a rare plant endemic to only four locations: one in Cave Springs, Georgia, one in Cherokee County, Alabama and two in Madison County, Tennessee. The species can grow up to three meters tall and has multiple showy yellow blooms making it a prime ornamental plant that is attractive to many different pollinators. This plant was designated as a federally endangered species in 2014 due to habitat loss. Currently, there is no recovery plan for *H. verticillatus*. Still, there are large gaps in knowledge related to basic biology of this plant and its importance in ecosystem services. In this study, microsatellite loci were utilized to investigate fine-scale population structure and clonal diversity of 206 *H. verticillatus* individuals found on two sampling sites within the Georgia population. Our results indicated the presence of two distinct genetic clusters that correlated with respective sampling site. However, admixture was present at the collection zones closest to the forested barrier separating the sites. Analyses of molecular variance indicated that the majority of variance (51%, $P<0.001$) was individually based, thus confirming high genetic differentiation ($F$st=0.20) and limited gene flow between *H. verticillatus* collection zones. The evidence of a population bottleneck in these sites suggests a recent reduction in population size that could possibly be due to habitat loss. In addition, high levels of linkage disequilibrium were found, implying that individuals within these sites are primarily reproducing asexually. Based on our results, although populations of *H. verticillatus* are limited and highly fragmented, they are still harboring moderate levels of genetic diversity ($H_{exp}$ = 0.50) and, in contrast to previous studies, high numbers of distinct genets. However, because of self-incompatibility and the ability to reproduce vegetatively, sexual reproduction is extremely limited in these populations. Results

presented here provided a better understanding of fine-scale genetic diversity and spatial distribution of *H. verticillatus* populations in Georgia. Combined with previous research findings, our results can underpin a novel recovery plan for *H. verticillatus* that could be utilized for conservation of this endangered species to promote its persistence in the wild.

# *Table of Contents*

1 Introduction ..... 1

2 Methods and Materials ..... 8
2.1 Sample Collection ..... 8
2.2 DNA Extraction ..... 9
2.3 Selection of Microsatellite Loci and Polymerase Chain Reaction ..... 9
2.4 Population Diversity ..... 11
2.5 Population Structure ..... 12
2.6 Demographic History ..... 13

3 Results ..... 14
3.1 Determining Polymorphic Loci ..... 14
3.2 Population Diversity ..... 14
3.3 Population Structure and Genetic Differentiation ..... 15
3.4 Demographic History ..... 17

4 Discussion ..... 18
4.1 Population Structure, Clonality and Genetic Diversity ..... 18
4.2 Conservation Implications and Further Research ..... 22

List of References ..... 28

Appendix ..... 40

Vita ..... 54

# *List of Tables*

Table 1. Fourteen microsatellite loci used to examine genetic diversity and spatial population structure of *Helianthus verticillatus* in Georgia and Tennessee, USA .....41

Table 2. Genetic diversity indices of *Helianthus verticillatus* samples from two sampling sites (C and R populations) analyzed as fourteen collection zones using fourteen microsatellite loci .....42

Table 3. Analysis of molecular variance (AMOVA) for *Helianthus verticillatus* across fourteen microsatellite loci for the all collection zones structured as one hierarchal group (A), two groups separated by sampling site (B), two clusters as indicated by STRUCTURE (C), and two clusters as indicated by discriminate analysis of principle components (DAPC) results (D).....43

Table 4. Pairwise population matrix of Nei's unbiased genetic distance of *Helianthus verticillatus* samples from two sampling sites (C and R) split into fourteen collection zones .....45

Table 5. Pairwise population differentiation for *Helianthus verticillatus* samples from two sampling sites (C and R) divided into fourteen subpopulations using 14 microsatellite loci. Fst values are based on 9,999 permutations.....46

Table 6. Bottleneck determination by sign tests for *Helianthus verticillatus* samples using fourteen microsatellite loci and grouped by STRUCTURE results (A), by sampling site (B) and by discriminate analysis of principle components (DAPC) results (C).....47

Table 7. Comparing sample size, genetic diversity and genetic structure from multiple species of *Helianthus* in past studies with data generated in this study on *H. verticillatus* .....48

## *List of Figures*

Figure 1. Diagram showing sampling (grid) methods employed at sites one (upper) and two (lower). Site one consisted of five contiguous 1 meter x 1 meter plots (only two of which are represented in the diagram). At site two, a 3 meters x 3 meters grid was created. The circles represent sampling locations within the grid. Sample names (RX.X) are inlayed within each circle and collection zones (R1-9) are differentiated by color. Empty circles represent locales where no plant was present and no sample was taken (lower diagram) .....49

Figure 2. STRUCTURE bar graphs representing genetic clusters (K=2-3) of samples from the two sampling sites (C and R) divided into fourteen collection zones. Each bar represents an individual sample and colors code membership of each sample's assigned cluster. Using Evanno's method, the presence of two genetic clusters (K=2) was found to be the best fit for this data.....50

Figure 3. Population structure and clustering for *Helianthus verticillatus* by collection zone using the Bayesian analysis program BAPS. The first graph (A) is assuming two clusters and the second graph (B) is assuming three clusters .. 51

Figure 4. Principle coordinates analysis (PCoA) using discriminant analysis of principle components (DAPC) method among fourteen collection zones of *Helianthus verticillatus* samples using fourteen microsatellite loci .....52

Figure 5. Minimum spanning network (MSN) of *Helianthus verticillatus* based on Bruvo's genetic distance for fourteen microsatellite loci. The nodes of the MSN represent individual multilocus genotypes (MLG) with the size and color representing population membership size and designated collection zone, respectively. To avoid overlapping nodes, the size was scaled to log1.75n, where n equals the node sample size. Lines between nodes represent genetic distance between MLG as determined by Prim's algorithm .....53

## 1. Introduction

*Helianthus verticillatus* (Small), commonly known as the whorled sunflower, is a herbaceous perennial endemic to four locations in the southeast United States (Matthews et al. 2002; Chafin 2010). The plants grow in large clonal clumps which can reach three meters in height (Matthews et al. 2002; Chafin 2010). They are unique among other plants in the *Helianthus* genus due to the characteristic whirling of hairy leaves around the stem (Matthews et al. 2002; Chafin 2010). The vigorous growth and showy yellow flowers of *H. verticillatus* may render the species a potential valuable ornamental plant and presumably a useful pollinizer in the wild as well as home garden, which is common for other *Helianthus* spp. (Schmidt et al. 1995; Chafin 2010).

*Helianthus verticillatus* was discovered in 1898 by Samuel McCutcheon, a faculty member at the University of Tennessee, near the town of Henderson in Chester County, TN (Seiler and Gulya 2004). Initially, labeled as *H. shcwenitzii* on one of the herbarium samples, it was officially annotated as *H. verticillatus* by Small (1898) (Matthews et al. 2002). After McCutcheon's original discovery, the plant was not observed in the wild again until 1994, where it was identified near the Coosa River in Floyd County, GA by James Allison (Matthews et al. 2002). Soon after, two more populations were discovered, the first in Cherokee County, Alabama in 1996 and the latter in Madison County, Tennessee in 1998 (Matthews et al. 2002; Ellis et al. 2008).

*Helianthus verticillatus* is morphologically similar to three other *Helianthus* spp., *H. angustifolius* (Narrow-leaved sunflower), *H. divaricatus* (Woodland sunflower), and *H. microcephalus* (Small headed sunflower), with shared traits of rounded hairy leaves and

small flower heads (Chafin 2010). In contrast to the three aforementioned species, *H. verticillatus* usually contains 3-4 leaves, which are arranged in a whorled, opposite pattern on the stem (Matthews et al. 2002; Chafin 2010). *Helianthus verticillatus* preferentially grows in prairie-like habitats, with open flood plains and wet depressions near the edges of forests (Chafin 2010).

This species was first speculated by Beatley in 1963 to be a hybrid of two other *Helianthus* species, *H. eggertii* and *H. angustifolius* (Beatley 1963; Matthews et al. 2002). However, this theory was impossible due to the difference in chromosome number between *H. eggertii* (n=51) and *H. angustifolius* (n=17) (Matthews et al. 2002). Due to similar morphology and overlapping habitats between *H. verticillatus* and other *Helianthus* spp. it was later speculated by Heiser *et al.* (1969) that *H. verticillatus* was a hybrid between *H. angustifolius* and *H. grosserratus*. Both species belong in the *Atroubens* section of the genus, have similar genetics (n=17), overlapping habitats, and *H. grosserratus* displayed verticillate leaf patterns in "exceptions" (Beatley 1963; Heiser et al. 1969; Matthews et al. 2002; Seiler and Gulya 2004; Ellis et al. 2006).

Ellis et al. (2006) examined the possibility of *H. verticillatus* being a hybrid between two other native species of sunflower, *H. angustifolius* and *H. grosseserratus*, while also investigating the genetic diversity and population of *H. verticillatus*. In this study 22 simple sequence repeats (SSRs), consisting of nuclear and chloroplast DNA, were used to assess genetic diversity in dense clusters of *H. verticillatus* plants (Ellis et al. 2006). Their results indicated moderate genetic diversity of *H. verticillatus* populations from three locations: one in Tennessee, Alabama and Georgia (Ellis et al. 2006). In addition, they observed significant

differences between populations of *H. verticillatus* and the proposed ascendant species (Ellis et al. 2006). It was thought that *H. verticillatus* could be a hybrid of other *Helianthus* species due to frequent hybridization within the genera, but did not descend from the specific cross of *H. grosseserratus* and *H. angustifolius* (Rieseberg 1991; Ellis et al. 2006). Their findings rejected the hypothesis that *H. verticillatus* was a hybrid, thus confirming that the plant was its own distinct species (Ellis et al. 2006). Further research focusing solely on this plant found high clonal diversity, despite predominate vegetative reproduction (Mandel 2010). Furthermore, there were far fewer distinct genetic individuals than previously thought (Ellis et al. 2006; Mandel 2010).

*Helianthus verticillatus* was classified as a federally endangered species in 2014 due to habitat loss by agricultural expansion and timber harvest (USFWS 2014). There are currently four known populations of *H. verticillatus*, two in Tennessee (TN), McNairy and Madison Counties, and two near the Georgia (GA)-Alabama (AL) border (Floyd County, GA and Cherokee County, AL), roughly 3.5 km apart (Ellis 2008; Mandel 2010). In Madison County, TN, the *H. verticillatus* population can be subdivided into two subpopulations less than 1.5 km apart; one very dense subpopulation located near Highway 45 and the other a sparse subpopulation located near railroad tracks and agricultural installation. The McNairy County, TN population is roughly 50 km south of Madison County and can be sectioned into two distinct populations along Prairie Branch Creek (Ellis et al. 2008). The GA and AL populations are located on timberland owned by Weyerhaeuser Corporation. The habitats of the GA and AL populations differ greatly from the TN populations, because they are wet prairies on undeveloped, scantily populated land (Ellis et al. 2008; Ellis and McCauley 2009; Mandel 2010). In addition, these

*H. verticillatus* populations are currently protected by the Nature Conservancy. The current management plan for these populations includes controlled burns on *H. verticillatus* inhabited land and frequent checks on population fitness.

Previous studies by Mandel (2010) and Ellis et al. (2006, 2008, 2009) have shown that *H. verticillatus* has moderate genetic diversity and drastic differences in fitness between populations. These studies also posit that *H. verticillatus* may be more prone to self-pollination and may have experienced a major decline in distinct genetic individuals in recent history (Mandel 2010). Without a concrete conservation plant for this species, these findings spurred the change from a low to high priority rank on the endangered species list and should instigate further efforts to protect this rare plant species (Mandel 2010).

Although some progress has been made regarding genetic diversity in *H. verticillatus* populations, our knowledge regarding basic biology, census data, and a lack of a well-defined conservation plan remain a major problem for preservation of this species (Ellis et al. 2006; Ellis 2008). The United States Fish and Wildlife Service (USFWS) with the United States Department of the Interior (USDOI) have designated the land on which this sunflower is endemic as a critical habitat (USFWS 2014). This designation only provides so much support for these plants because, as outlined by the Endangered Species Act (ESA), critical habitats have little effect on the land if federal funds are not involved (ESA 1973). Because of this, corporate and private landowners (i.e. The Nature Conservancy and the Weyerhaeuser Corporation in the case of the Alabama/Georgia populations) have complete control over protection of this plant, including but not limited to provision of funds for habitat maintenance as well as best management

practices to landowners for this species. However, there is currently no protective funding for the Tennessee populations, leaving those populations completely vulnerable.

To protect and conserve the biodiversity of *H. verticillatus*, we have to provide a better understanding of its biology, genetics and overall population structure. Measured by the number of unique life forms, their equitability, genetic variability, and biodiversity dictates processes such as ecosystem services and the ability of that system to function (Cardinale et al. 2012). By protecting the diversity of the ecosystems in which this species inhabit, we maintain its equilibrium and therefore, ecosystem's ability to function (Cardinale et al. 2012). To better understand how to enhance conservation efforts for *H. verticillatus*, this study utilized microsatellite loci to determine genetic variation and spatial dynamics of the few remaining small populations of this plant while trying to discern the effect of high levels of clonality within them (Pashley et al. 2006).

In the past few decades, population genetics has become an important tool in conservation biology (Ellis 2008). Analyses in the realm of population genetics allow for the mapping of genetic variation, as well as the assessment of gene flow and effective population size of the species in question (Wyatt 1992). These techniques are also effective in determining the effect of geographic subdivision of populations in rare and endangered species (Wyatt 1992). While these approaches are effective at providing insight into the genetics between populations, they can also be applied to the fine-scale genetics within populations (Ellstrand and Roose 1987; Wyatt 1992; Chung and Epperson 2000). For the rare plant species such as *H. verticillatus,* genetic diversity tends to be diminished due to small population sizes that can often be separated by hundreds of kilometers (Willi et al. 2006). This often results in reduced

levels of fitness and lower heterozygosity within the populations (Hamrick and Godt 1990; Willi et al. 2006; Ellis 2008). With the lack of knowledge concerning *H. verticillatus*, including no census or historical range records, the use of population genetics, especially at a fine-scale, could give more insight into the possible history and biology of this plant (Maruyama and Fuerst 1985; Hamrick and Godt 1990; Matthews et al. 2002; Ellis 2008). Understanding the current state of spatial dynamics and genetic diversity of *H. vertcillatus* coupled with experimental data of historical genetics could prove to be paramount for conservation of this species, as it has with other species in similar predicaments (Brzosko et al. 2002; Walck et al. 2002; Willi et al. 2006; Bowen 2011).

Currently there is limited knowledge regarding fine-scale population structure and genetic diversity of *H. verticillatus*. In previous studies performed on a large geographic scale, it was found that while the populations are isolated and mainly clonal, there is high genetic diversity found within the populations (Ellis et al. 2006; Ellis and McCauley 2009; Mandel 2010). These studies also found high clonal diversity, the presence of many polymorphic genotypes in a species demonstrating predominantly asexual reproduction, in the respective populations (Ellstrand and Roose 1987; Mandel 2010). Focusing on this plant at a smaller geographical scale could give new insights into the recent history, possible hybridization, mating systems and further refine existing conservation plans for *H. verticillatus* (Ellis 2008). The main objective of this study was to assess clonal diversity and spatial distribution of *H. verticillatus* at two geographically close populations in Georgia. Based on the biology and previous research of *H. verticillatus*, our hypothesis is that there will be diminished numbers of distinct genets within the sites sampled. However, we also hypothesize that these two sites will harbor high genetic

diversity and should be spatially structured into discrete genetic clusters. With expanded knowledge of the fine scale genetic diversity and population structure of these two populations of *H. verticillatus*, augmenting current plans and the creation of a solid recovery plan will be much more feasible.

## 2. Materials and Methods

*Sample Collection*

*Helianthus verticillatus* leaf samples (n=206) were collected from two separate sites located near Cave Spring, GA. Site one consisted of five contiguous 1 x 1 m quadrants of *H. verticillatus*. All of the stems were counted in each quadrant and five leaves per stem were taken at random from no more than 30% of the individual stems. Site one yielded a total of 74 samples, 27 from the first quadrant, 17 from the second, 5 from the third, 11 from the fourth and 14 from the fifth quadrant (Fig. 1). Site two was less than 1 km south east of site one in a large clearing spanning about 60 x 30 m. This plot was divided into a 3 x 3 m grid (Fig. 1) with five leaf samples taken from one individual stem at the intersects of each quadrant and from one individual stem at the center of the quadrant (Escaravage et al. 1998) (Fig. 1). This sampling method was adapted from Escaravage et al. (1998) due to the clonal nature of this plant in an effort to only sample a single genet at each intersect. Some intersects did not contain *H. verticillatus* specimens and were excluded from this study. The second site yielded a total of 132 samples.

The sampling methods at both sites were different for multiple reasons. First of all, there were two sites to check for spatial structure of these populations. With the forested barrier, the individual sampling sites, in theory, should cluster together. At the first site, the 1 x 1 m quadrants were contiguous to assess the potential clonal spread observed in previous studies (Lienert 2004; Ellis et al. 2006; Ellis 2008; Ellis et al. 2008; Ellis and McCauley 2009; Mandel 2010). At site two, the sampling locales (grid intersects) were separated by three meters with the exception of the center-grid (~1.5 m from the intersects) locations. This

method was employed to avoid collecting genetically identical individuals, per previous studies, and remove potential bias caused by clonality from this portion of the study (Lienert 2004; Ellis et al. 2006; Ellis 2008; Ellis et al. 2008; Ellis and McCauley 2009; Mandel 2010). All samples (n = 206) were placed in plastic bags containing silica gel at a ~10:1 ratio (silica gel to plant material weight) and stored on ice at the time of sampling to prevent degradation (Chase and Hills 1991). These samples were then stored at -80°C until DNA extractions were completed.

*DNA Extraction*

Portions of a single leaf from each sample were placed into sterile 2 ml conical screw cap microcentrifuge tubes (Fisherbrand, Pittsburg, PA, USA) with sterile 13 mm zirconia/silica beads (BioSpec Products, OK, USA) and submerged into liquid nitrogen for 2 min. Samples were then homogenized twice using a Bead Mill 24 (Fisher Scientific, Walther, MA, USA) for 20 sec each with a 5 min period in liquid nitrogen between each homogenization step. DNA extraction was completed using a modified protocol of the DNeasy Plant Mini kit (Qiagen, Valencia, CA, USA) including 2% by volume polyvinylpyrrolidone (PVP) and 4µl of RNaseA added to the lysis buffer. After addition of P3 buffer, samples were incubated at -20°C until frozen. Genomic DNA was quantified using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). DNA was stored at -20°C until needed for amplification.

*Selection of Microsatellite Loci and Polymerase Chain Reaction*

Previously published microsatellite loci developed from *H. annuus* and found not to be species specific (Ellis et al. 2006; Pashley et al. 2006; Ellis 2008; Mandel 2010) were screened for

the presence of polymorphisms, amplification and consistency among the collected samples of *H. verticillatus*. Initial testing for selection of microsatellite loci used four *H. verticillatus* leaf samples from both Tennessee locations and from the Georgia location. From 48 microsatellite loci, 15 tri- and tetra-repeat microsatellites were selected and used in this study. Polymerase chain reaction (PCR) was completed in 10 µl reactions with 10 ng/µl genomic DNA (gDNA), 2.5 µM of both forward and reverse primer, 0.5 µl dimethyl sulfoxide (DMSO) (Thermo Fisher Scientific, Waltham, MA, USA), 4 µl (0.8X) GoTaq Colorless Mastermix (Promega, Madison, WI, USA), and brought to the final volume with sterile Nanopure water (Thermo Fisher Scientific). Reactions were performed using a Mastercylcer Pro Automatic Thermal Cycler (Eppendorf Biotech Company, Hamburg, Germany) at the following conditions: 95°C for 3 min, followed by 10 cycles of 94°C for 30 sec, 65°C lowering 1°C per cycle to a final 55°C for 30 sec, then 72°C for 45 sec, another 30 cycles of 94°C for 30 sec, 55°C for 30 sec, 72°C for 45 seconds and a final elongation step at 72°C for 20 min (Ellis 2008). Resulting amplicons were then analyzed on the QIAxcel Capillary Electrophoresis system (Qiagen, Valencia, CA, USA) using a 15/600 base pair (bp) internal marker and scored with a 25 bp DNA size marker to assess raw allele length (Wang et al. 2009; Dean et al. 2013). A positive control, a sample that amplified consistently across all primers, was used in every 96-well plate as well as a negative control of sterile Nanopure water (Thermo Fisher Scientific) to check for consistency of results. If either the positive control failed to amplify, or the negative control amplified in any of the plates, the whole set of reactions was repeated.

*Population Diversity*

FLEXIBIN v2 (Amos et al. 2007) was used to bin the raw allelic data into allele classes and to remove monomorphic loci. The resulting data set was used for all further analyses. Samples in this study (n=206) were then grouped into 14 collection zones: 5 representing the 1 x 1 m quadrants sampled at the first site and 9 for each row of the grid at the second site (Fig. 1). All binned data were clone corrected using POPPR v2.1.1 (Kamvar et al. 2014) to remove any identical multilocus genotypes (MLG) from each collection zone. Clone correction was used to avoid any biases that could be caused by identical MLGs in further analyses.

All subsequent data analyses were analyzed using the RStudio environment (RStudio 2012; RCoreTeam 2016). Multiple packages for R were used for data analyses in this study. The package POPPR v2.1.1 was used to calculate various genetic diversity indices including the index of association (Ia), which takes into account both allelic richness and evenness ($E_5$) of a collection zone (Hill 1973; Shannon 2001). In addition, POPPR was also used to calculate the standard index of association ($\overline{r}_d$), a measure of linkage disequilibrium (Agapow and Burt 2001) using 10,000 permutations for each collection zone. The number of private alleles, alleles found only in one collection zone (Szpiech and Rosenberg 2011), and Nei's genotypic diversity ($H_{exp}$) (Nei 1978) were also analyzed with POPPR. The package hierfstat v0.04-22 (Goudet 2005) was used to calculate the pairwise population differentiation and the package adegenet v2.1.1 (Jombart 2008) was utilized to calculate Nei's pairwise genetic distance.

*Population Structure*

Clustering and population structure of *H. verticillatus* were assessed with STRUCTURE v2.3.4 (Pritchard et al. 2000) using a Bayesian Monte Carlo Markov Chain (MCMC) method. The parameters used in STRUCTURE included a burn-in period of 500,000 with 500,000 MCMC repetitions of 30 iterations at K=1-10. STRUCTURE HARVESTER web v0.6.94 (Earl 2012) was used to infer the optimum K value, using Evanno's method (Evanno et al. 2005) to represent the most probable number of genetic clusters. POPHELPER web v1.0.10 (Francis 2016) was utilized to visualize the optimum value of K from the previous analyses. BAPS v5.0 (Corander et al. 2008) was also used to infer overall population structure of the sample set. This program allows users to specify the number of genetic clusters to be tested, depending on the hypotheses and research question(s) of interest (Corander et al. 2008). In this study, we used BAPS as set to test for the presence of two clusters among *H. verticillatus* individuals under the assumption that both collection sites would group into two distinct populations due to high levels of clonal propagation (Ellis et al. 2006; Ellis 2008; Ellis and McCauley 2009; Mandel 2010). The R package PopGenReport v3.0.0 (Adamack and Gruber 2014) was used to visualize a principle components analysis (PCoA) using the discriminant analysis of principle components (DAPC) method (Jombart et al. 2010).

Genetic differentiation was calculated using an analysis of molecular variance (AMOVA) with Arlequin v3.5.2.2 (Excoffier and Lischer 2010). To assess genetic differentiation of *H. verticillatus* samples in this study, four variance partitions were used: one in which all collection zones were analyzed as a single hierarchical group, another where the collection zones were analyzed by collection site, one which was analyzed with STRUCTURE results and a final one

analyzed according to DAPC assignment. Bruvo's distance, which estimates genetic distance between individuals rather than between collection zones (Bruvo et al. 2004), was calculated using POPPR and subsequently applied to create a minimum spanning network (MSN).

*Demographic History*

The program BOTTLENECK v1.2.02 (Cornuet and Luikart 1996) was used to determine whether there had been any recent bottleneck or expansion among *H. verticillatus* collection zones. Sign and Wilcoxon tests were employed to establish whether the loci used in this study were in the mutation-drift equilibrium (Cornuet and Luikart 1996). The Sign test posits a null hypothesis of differences between observed and expected heterozygosity, while the Wilcoxon test assumes a null hypothesis of no significant excess of heterozygosity (Luikart et al. 1998; Piry et al. 1999). Three mutation models were used with 10,000 iterations each: infinite allele model (I.A.A.), stepwise mutation model (S.M.M.), and the two-phase model (T.P.M.) at default settings (Piry et al. 1999). The data was analyzed grouping samples by site as well as according to STRUCTURE and DAPC results.

## 3. Results

*Determining Polymorphic Loci*

Screening for suitable loci indicated that fourteen tri- and tetra-repeat microsatellites out of 48 primer pairs were polymorphic and amplified in all our tested samples (Table 1). The other 34 loci were rejected based on weak or no reactions, allele sizes out of predicted ranges, or the lack of polymorphisms. These microsatellite loci were selected for further population assessments and binned prior to data analyses. Binning the lengths of the amplicons produced 2-16 different allelic classes for each of the fourteen loci (Table 1). The amplicon size difference of these primers varied greatly over these classes, ranging from 4 (HV41) to 66 (HV42) base pair differences.

*Population Diversity*

Clone correction of the original dataset (n=206) removed three samples from subsequent analyses. Samples removed through clone correction were from three separate collection zones (C1, C5, and R7). The clone corrected data (n=203) was used for all subsequent analyses with 14 polymorphic loci. *Helianthus verticillatus* individuals were divided into 14 different collection zones based on location and quadrant in which samples were collected (C1-C5 corresponding to the first collection site where five different quadrants were sampled; R1-R9 corresponding to site two and the nine different rows from the sampling grid).

The Shannon-Weiner index of multilocus genotype (MLG) diversity (H) ranged from 1.61 in collection zone C3 to 3.26 in collection zone C1, with an overall average of 2.58 (Table 2). Evenness ($E_5$) was 1 for all collection zones and 21 private alleles were found across 8 of the

collection zones (Table 2). Only one collection zone from site one (C2) had private alleles (n=2). In site two, the majority of private alleles were found in all but two of the collection zones (R6 and R7) with the most being found in collection zone R9 (n=7) (Table 2). Nei's gene diversity ($H_{exp}$) ranged from 0.36 (collection zone C5) to 0.59 (collection zone R8) with an average of 0.50. However, site one had lower gene diversity (0.44) compared to site two (0.53) on average. The standardized index of association ($\overline{r}_d$), a measure of linkage disequilibrium, had an average value of 0.13 with a maximum value of 0.28 (R7) and was significant for all collection zones with the exception of zone R6 (*P*=0.07) (Table 2).

*Population Structure and Genetic Differentiation*

*Helianthus verticillatus* individuals grouped into two distinct genetic clusters (ΔK=2), based on the STRUCTURE result, utilizing Evanno's method (Evanno et al. 2005) within STRUCTURE HARVESTER (Earl 2012). STRUCTURE results were subsequently visualized using the POPHELPER v (Francis 2016) platform (Fig. 2). When the *H. verticillatus* samples were grouped into two clusters, (ΔK=2), individuals from the R1 collection zone indicated the presence of admixture, which was very limited within other *H*. *verticillatus* collection zones (Fig. 2). Though Evanno's method indicated ΔK=2 as the most probable result, data was also visualized as ΔK=3 (Fig. 2). Clustering patterns for ΔK=3 indicated similar levels of admixture for the R1 collection zone, but additional admixture across other collection zones (Fig. 2). BAPS produced similar results to STRUCTURE when analyzed as 2 clusters (K=2) and 3 clusters (K=3) (Fig. 3). However, these analyses did differ slightly in that BAPS R1 was clustered consistently with the C zones rather than its respective sampling site (Fig. 3).

The analysis of molecular variance (AMOVA) of the collection zones as one hierarchal group indicated that much of the variation was found within the individuals (76.04%, *P*<0.001; *F*st=0.24) (Table 3). When data was analyzed with the collection zones as two distinct groups (C and R sampling sites), the majority of the variation was found within individuals (50.96%, *P*<0.001), rather than among individuals within collection zones (30.17%, *P*<0.001), and among collection zones (18.87%, *P*<0.001) (Table 3). When data was partitioned based on STRUCTURE results the lowest variation was found among two the collection zones (20.18%, *P*<0.001; *F*st of 0.20), compared to among individuals within two genetic clusters (28.92%, *P*<0.001), and within individuals (51%, *P*<0.001) (Table 3). When samples were analyzed grouped by DAPC designation the results were similar to those of the AMOVA when grouped by sampling site and STRUCTURE, with the majority of variation occurring within individuals (50.72%, *P*<0.001) (Table 3).

Nei's pairwise genetic distance was lowest within the first sampling site, particularly between collection zones C4 and C5 (Table 4). The highest genetic distance was found between collection zones C5 of the first sampling site and R9 of the second site (Table 4). Pairwise population differentiation ($F_{st}$) corresponded with Nei's genetic distance, with the highest amount of differentiation between collection zones C5 and R9, and the lowest between collection zones C4 and C5 (Table 5). Principle coordinates analysis (PCoA) matched the STRUCTURE and BAPS results clustering into two main groups, one predominately site one and site two with admixture at collection zone R1 (Fig. 4). The MSN corresponded with BAPS, STRUCTURE and PCoA results, showing two distinct clusters with MLGs from zone R1 grouping closer to MLGs from the first sampling site (C1-C5) (Fig. 5). Genetic distance between MLGs in

the MSN is represented by the thickness of the line between each MLG node (Kamvar et al. 2015). The size and color of each node represents the number of samples and collection zone representing each MLG respectively (Kamvar et al. 2015).

*Demographic History*

The program BOTTLENECK indicated the presence of a recent population bottleneck for all three grouping analyses. The Sign, Wilcoxon, and the standardized difference tests across three mutation models showed a significant heterozygote excess when the collection zones were grouped by sampling site, with both shifting from the mutation equilibrium model (Table 6). When the collection zones were grouped by STRUCTURE and DAPC assignments, the same results were observed (Table 6).

## 4. Discussion

*Population Structure, Clonality and Genetic Diversity*

Our results indicated moderate genetic diversity, high levels of clonality, the presence of spatial structure, and evidence of recent bottleneck among *H. verticillatus* individuals, thus supporting our hypothesis of clonality and population structure. Both sampling sites harbored moderate levels of genetic diversity and in contrast to our proposed hypothesis, a high number of distinct genets (N = 203 of 206 total samples). In all but two collection zones we found that each sample was a unique multilocus genotype (MLG). The moderate diversity observed within these collection sites appear to contradict much of the current research concerning rare or endangered plant species (Ellstrand and Elam 1993; Gitzendanner and Soltis 2000; Willi et al. 2006). However, we did not find evidence of the low or high genetic diversity reported in earlier studies completed by Mandel et al. (2013) and Gevaert et al. (2013) respectively concerning two other endangered species of *Helianthus* in the United States.

In their recent study, Mandel et al. (2013) investigated genetic diversity of the rare and endangered sunflower *H. niveus* spp. *tephrodes* (Algodones sunflower), which is native to Southern California and Mexico. The authors found lower levels of genetic diversity ($H_{exp}$ = 0.31) in *H. niveus* spp. *tephrodes* when compared to other rare (*H. verticillatus;* $H_{exp}$ = 0.48) and endemic (*H. porteri* (Porter's Sunflower); $H_{exp}$ =0 .69) *Helianthus* spp. (Table 7) (Gevaert et al. 2013; Mandel et al. 2013). *Helianthus niveus* spp. *tephrodes* showed a similar genetic structure in which clustering was correlated with geography, but the sampling sites were considerably further apart than in the our study (Mandel et al. 2013). However, Gevaert et al. (2013) study focused on *H. porteri* had a larger number of individuals (n=200) compared to *H. niveus* ssp.

*tephrodes* (n=119) and *H. verticillatus* (n=71) studies, which could result in biasing the outcomes of genetic diversity (Table 7). Because of *H. porteri's* declining habitat (rocky outcroppings in the southeast United States), there have been reintroduction efforts and this study compared these to native populations (Gevaert et al. 2013). Although their findings indicated high genetic diversity, they found a lack of population structure.

Both of the previous studies compared their target species with previous research by Mandel (2010) concerning *H. verticillatus*. Here, we found similar levels of genetic diversity as reported in previous findings, although their study included fewer individuals from a much larger geographic distribution (N = 71; TN, AL and GA sites) (Mandel 2010). However, these studies found much higher and lower levels of diversity in *H. porteri* and *H. niveus* spp. *tephrodes* respectively (Gevaert et al. 2013; Mandel et al. 2013). The lower number of samples of *H. verticillatus* could be causing biases in statistical analyses, as with more data the more complete the representation of a study species is (Nei and Roychoudhury 1974; Nei 1978). Nei and Roychoudhury (1974) speculated that in leu of high sample numbers, an increased number of loci could be tested. This could also play a part in the drastic differences in species diversity as the number of loci tested in these studies varied from 11 to 18 (Table 7) (Nei and Roychoudhury 1974; Nei 1978; Mandel 2010; Gevaert et al. 2013; Mandel et al. 2013).

Habitat fragmentation and small population size can increase inbreeding among individuals thus limiting gene flow, which can reduce genetic variation and population fitness (Young et al. 1996; Storfer 1999; Willi et al. 2006). The presumed high levels of vegetative propagation in *H. verticillatus* could provide an explanation for reduced diversity within these populations (Ellstrand and Roose 1987; Ellis et al. 2006; Mandel 2010). However, the moderate

diversity found in our study may be explained by the possibility that sexual reproduction was more frequent within these sampling sites in the past and the harbored diversity is purely a relic (Esselman et al. 1999), or that sexual reproduction is occurring but resulting in few viable seeds (Ellis 2008).

Widén and Anderson (1993) indicated that moderate to high genetic diversity has been found in plants in scenarios similar to that of *H. verticillatus*. Studies involving *Senecio integrifolius*, an endangered plant found in Sweden, showed that spatial population structure may play a large role in genetic diversity (Ellstrand and Elam 1993; Widén and Andersson 1993). The target species in this study exhibited high levels of diversity despite the samples coming from small fragmented populations (Widén and Andersson 1993). The authors proposed that small populations shelter genetic diversity and when habitat destruction causes population fragmentation, many generations are required to limit variations within the populations (Widén and Andersson 1993). *Helianthus verticillatus,* which is endemic to these sampling sites and a possible relic of a larger population may contribute to the moderate levels of diversity observed in our study (Ellstrand and Elam 1993). The high levels of spatial and genetic structure found in our sampling sites could be contributing to the moderate level of diversity as well.

Linkage disequilibrium has become a very important tool in determining whether a sampled species is reproducing asexually versus sexually (Flint-Garcia et al. 2003; Slatkin 2008). Linkage disequilibrium is the nonrandom association of alleles occurring in at least two loci and can be used to understand mutations in populations as well as the effects of natural selection. However the use of linkage disequilibrium in studies pertaining to plants is limited (Flint-Garcia et al. 2003; Slatkin 2008). We found significant linkage disequilibrium for each collection zone

indicating that *H. verticillatus* is more likely asexually reproduced, supporting our hypothesis of predominately vegetative reproduction (Tian et al. 2015; Kamvar et al. 2017). The evidence of a recent bottleneck reinforces the presence of high linkage disequilibrium (Flint-Garcia et al. 2003). This is caused by genetic drift and the low number of allelic combinations passed to future generations (Flint-Garcia et al. 2003). Sugarcane (*Saccharum* spp.) is an example of high linkage disequilibrium in plant species which, like *H. verticillatus*, has been subject to a genetic bottleneck and is propagated mainly through asexual means (Flint-Garcia et al. 2003).

In accordance with our hypothesis, genetic structure was based upon the sampling locations. However, there is some admixture observed between the two sites, with the majority present in collection zone R1. One plausible explanation is that the collection zone (R1) is located on the forested barrier between site two (R1-R9) and site one (C1-C5). With the close proximity (less than 1km) of these two sites, it is possible that pollinators harbored in the forested barrier could be transferring pollen between the sites. It would be presumed that native and honey bees would be the primary pollinator for this as well as other sunflowers, but very little research has been done on pollinators in wild *Helianthus* species (Heiser et al. 1969). Some bee species travel only to neighboring plants and have a flight distance of no more than 3m (Schmitt 1980). However, it has been observed that physical barriers and geographic distance between populations tends to keep pollinators within individual populations (Schmitt 1980; Loveless and Hamrick 1984).

Habitat fragmentation almost always leads to a reduction in population numbers, which inevitably leads to a reduction in diversity resulting in a population bottleneck (Amos and Harwood 1998). Our results showed significant heterozygote excesses when analyzed by DAPC

and STRUCTURE results as well as by sampling location, indicating population bottleneck caused by a reduction in population numbers (Barrett 1991). Because STRUCTURE and BAPS analyses assume that the loci studied are in Hardy-Weinberg equilibrium (HWE), the evidence of heterozygote excesses could contradict those results (Pritchard et al. 2000; Corander et al. 2008; Hubisz et al. 2009). However, when analyzed by parameters set by DAPC and Bruvo's distance, neither of which assume the presence of HWE, the results pointed to two genetic clusters in accordance with STRUCTURE (Fig. 3; Fig. 4; Fig. 5) (Bruvo et al. 2004; Jombart et al. 2010). Further supporting the conclusion of two genetic clusters based off sampling site, Nei's genetic distance and $F_{st}$ results suggest the presence of two distinct clusters (Table 4; Table 5).

The moderate levels of diversity ($H_{exp}$ = 0.50) in these collection zones suggests that the bottleneck occurred recently and a reduction in variation has not been able take place (Barrett 1991). Likewise, clonal reproduction may have led to an increase in heterozygosity by way of somatic mutations over time like the apomatic fern *Dryopteris remota* (Wood fern) or tree species like *Populus tremuloides* (Trembling aspen) (Antolin and Strobeck 1985; Judson and Normark 1996; Schneller et al. 1998; Ally et al. 2008). Because of this reduction in the number of individuals and the spatial structure of these populations, they are at an increased risk for inbreeding depression and a loss of overall fitness (Amos and Balmford 2001).

*Conservation Implications and Future Research*

There are number of successful conservation efforts such as *Potentilla robbinsiana* (Robins' Cinquefoil) and *Echinacea tenneseensis* (Tennessee Purple Coneflower) that could be used as a model system for a recovery plan and preservation of biodiversity in *H. verticillatus*

fragmented populations. *Potentilla robbinsiana* was listed as endangered in 1980 and like our study species was restricted to small, fragmented populations in the White Mountains of New Hampshire (USFWS 2002). A recovery plan was finalized in 1991 that included screen walls with alert signs to separate it from nearby walking trails, educational posters, transplanting into native habitats and even trail relocation (USFWS 2002). The U.S. Fisheries and Wildlife service also created parameters for "viable populations" that would consist of 50 or more plants per population (USFWS 2002). Subsequently, Robins' Cinquefoil was removed from the Endangered Species List in 2002 (USFWS 2002).

*Echinacea tennesseensis* was officially listed as endangered in 1976 due to habitat loss and recreational development (USFWS 2011). Like both *H. verticillatus* and *Potentilla robbinsiana*, this species was also found in small fragmented populations in the Tennessee counties of Davidson, Rutherford and Wilson (USFWS 2011). The recovery plan for this species was completed in 1989 and defined one condition to be delisted: the securement of five colonies (populations), three of which are self-sustaining (USFWS 2011). The definition of a "colony" was at least 15% flower cover of 669 square meter plot of suitable habitat (USFWS 2011). *Echinacea tennesseensis* was also part of a seed storage campaign, placed in multiple botanical gardens and sold commercially (Walck et al. 2002). It was also the subject of experiments exploring its genetics in relation to conservation efforts, *in vitro* propagation methods and seed viability (Baskauf et al. 1994; Walck et al. 2002; Kim et al. 2004; Sauve et al. 2004).

A close relative of and a plant *H. verticillatus* was previously mistaken for, *Helianthus schwenitzii* (Schwenitzii's Sunflower), has also been declared federally endangered (USFWS

1991; Grubbs and Wynes 2015). Conservation efforts are ongoing in its endemic habitats in North and South Carolina (USFWS 1991; Grubbs and Wynes 2015). Like our study species, the reason for *H. schwenitzii's* endangered status is due to change in land management, removing naturally occurring wild fires and grazing herbivores (Grubbs and Wynes 2015). *Helianthus schwentzii* also has a limited conservation plan with few concrete objectives and described critical habitat (USFWS 1991). In their study, Grubbs and Wynes (2015) focused *H. schwenitzii's* reproduction with a focus on seed germination and vegetative reproduction from its tubers. In one experiment the authors utilized gibberellic acid treatments (1000 mg/L) to improve germination rates from 15% (untreated) to 40% (treated), increasing their potential survivability (Grubbs and Wynes 2015).

Because *H. schwenitzii* is endemic to lands prone to wildfires they also tested seed germination in a charred soil and organic matter mix. They found the seeds had a 5% germination rate in the soil mixture that was not burned, but had 7% and 16% germination rate in mixes that were burned for 15 and 30 minutes, respectively (Grubbs and Wynes 2015). Fire could play an important role in *H. verticillatus* as it was demonstrated in germination experiments with *H. schenitzii* (Grubbs and Wynes 2015). With increases in population, timber harvesting and agricultural operations, there has been a significant decrease in naturally occurring wildfires due to technological advances and prescribed burning (Whelan 2002; Fernandes and Botelho 2003). However it has been demonstrated many times over that in prairies, such as the native habitat of *H. verticillatus*, fire can be extremely beneficial (Groeschl et al. 1992; Whelan 2002; Fernandes and Botelho 2003; Hochkirch and Adorf 2007). Groschl et

al. (1992) observed an increase in species diversity after prescribed burns, thought to be because of the removal of mid-canopy growth.

The fragmented populations and potential habitat loss found in *H. verticillatus* could be an outcome of reduction in naturally occurring wildfires and controlled prescribed burning. As seen in *H. schwenitzii*, wildfire have decreased but haven't been have not been completely eradicated in the areas inhabited by *H. verticillatus*. This is due in part to prescribed burns being employed for hazard reduction as well as conservation purposes (Whelan 2002; Fernandes and Botelho 2003). Because these controlled burns are less intense than naturally occurring wildfires, there is less mortality of upper and mid-canopy plants, thus removing the benefit for smaller herbaceous species needing open spaces for colonization..

In the realm of species preservation, little is known and planned for *Helianthus verticillatus* conservation efforts. In addition, our knowledge regarding pathogens and insects affecting this species is limited. In April 2015, powdery mildew caused by *Golovinomyces ambrosiae* was observed on *H. verticillatus* at two residential sites in Knoxville, TN (Trigiano et al. 2015), although the pathogen has not been confirmed to affect wild populations of *H. verticillatus* (Trigiano et al. 2015). Later that year in October, small necrotic lesions were observed on wild plants at the Cave Springs, GA as well as the Alabama sites. Our investigations confirmed that the causal pathogen was *Alternaria alternata* (Edwards et al. 2016). However, *A. alternata* appeared to cause only cosmetic damage as the lesions and premature loss of foliage was non-fatal when observed on greenhouse propagated plants (Edwards unpublished). No other pathogens have been observed or confirmed on this plant. However, other *Helianthus*

species are known to be affected by at least 30 pathogens, some of which are fatal to infected plants, that could be problematic in the future for *H. verticillatus* (Markell et al. 2015).

There are many avenues to frame potential conservation efforts and ample opportunities for continued research, however there is no federally mandated conservation plan for *H. verticillatus*. The similarities of these former cases could provide insight into development of a plan for *H. verticillatus* and provide assistance in creation of a recovery plan. This and previous studies have revealed valuable insights into the genetics contained within these populations (Ellis et al. 2006; Ellis 2008; Ellis et al. 2008; Ellis and McCauley 2009; Mandel 2010). The presence of moderate genetic diversity within these populations could prove useful in translocation of specimens between populations to promote gene flow (Keller et al. 2001). However, many studies have pointed to translocations being a double-edged sword (Lienert 2004). On one hand, the introduction of different genetics could increase diversity and lower the risk of inbreeding depression (Lienert 2004). On the other, these populations may be specifically adapted to that particular habitat and the introduction of new alleles could lower the fitness of highly adapted plants by way of breaking up beneficial gene complexes (Lienert 2004).

However, there is still much that is not understood about this *H. verticillatus'* biology and its relation to the environment. Further research into seed viability and *in vitro* propagation could prove to be crucial for preservation of this species in seed banks and botanic gardens (O'Donnell and Sharrock 2017; Volis 2017). The continued monitoring of these populations will provide much needed data of the differences of fitness from site to site and how this species interacts with its environment. With this and previous research coupled with an expanded

understanding of *H. verticillatus's* predicament and ecology, a strong recovery plan could promote this rare plant's persistence in its native habitat (USFWS 2002; Bowen 2011).

Combining previous research on *H. verticillatus,* other *Helianthus* species, data on endangered and recovered plants, we recommend the five following actions: 1) Continued research of *H. verticillatus's* biology and ecology; 2) placement of this species in botanical gardens and seed banks; 3) exploration for unknown populations; 4) the continued monitoring of known populations; 5) *in vitro* propagation and germplasm preservation.

## List of References

Adamack AT, Gruber B. 2014. PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution* **5**: 384-387.

Agapow PM, Burt A. 2001. Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* **1**: 101-102.

Ally D, Ritland K, Otto S. 2008. Can clone size serve as a proxy for clone age? An exploration using microsatellite divergence in *Populus tremuloides*. *Molecular Ecology* **17**: 4897-4911.

Amos W, Balmford A. 2001. When does conservation genetics matter? *Heredity* **87**: 257.

Amos W, Harwood J. 1998. Factors affecting levels of genetic diversity in natural populations. *Philosophical Transactions of the Royal Society B: Biological Sciences* **353**: 177-186.

Amos W, Hoffman J, Frodsham A, Zhang L, Best S, Hill A. 2007. Automated binning of microsatellite alleles: problems and solutions. *Molecular Ecology Notes* **7**: 10-14.

Antolin MF, Strobeck C. 1985. The population genetics of somatic mutation in plants. *The American Naturalist* **126**: 52-62.

Barrett SCH. 1991. Genetic and evolutionary consequences of small population size in plants: implications for conservation. *Genetics and Conservation of Rare Plants*: 3-30. Falk DA and Holsinger KE, eds. Oxford University Press.

Baskauf CJ, McCauley DE, Eickmeier WG. 1994. Genetic analysis of a rare and a widespread species of *Echinacea* (Asteraceae). *Evolution* **48**: 180-188.

Beatley JC. 1963. The sunflowers (genus *Helianthus)* in Tennessee. *Journal of the Tennessee Academy of Sciences* **38**: 135-154.

Bowen B. 2011. Natural areas protection at its best; protecting the Tennessee purple coneflower (*Echinacea tennesseensis*). *Natural Areas Journal* **31**: 326-330.

Bruvo R, Michiels NK, D'Sousza TG, Schulenburg H. 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* **13**: 2101-2106.

Brzosko E, Wróblewska A, Ratkiewicz M. 2002. Spatial genetic structure and clonal diversity of island populations of lady's slipper (*Cypripedium calceolus*) from the Biebrza National Park (northeast Poland). *Molecular Ecology* **11**: 2499-2509.

Cardinale BJ, Duffy JE, Gonzalez A, Hooper DU, Perrings C, Venail P, Narwani A, Mace GM, Tilman D, Wardle DA et al. 2012. Biodiversity loss and its impact on humanity. *Nature* **486**: 59-67.

Chafin LG, Owers, K. 2010. Species account of *Helianthus verticillatus*. Georgia Wildlife. https://georgiawildlife.com/sites/default/files/wrd/pdf/.../whorled_sunflower_2016.pdf

Chase MW, Hills HH. 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. *Taxon* **40**: 215-220.

Chung M, Epperson B. 2000. Clonal and spatial genetic structure in *Eurya emarginata* (Theaceae). *Heredity* **84**: 170-177.

Corander J, Marttinen P, Sirén J, Tang J. 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics* **9**: 539-553.

Cornuet JM, Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* **144**: 2001-2014.

Dean DA, Wadl PA, Hadziabdic D, Wang X, Trigiano RN. 2013. Analyzing microsatellites using the QIAxcel system. In: Kantartzi S. (eds) Microsatellites. Methods in Molecular Biology (Methods and Protocols), vol 1006. Humana Press, Totowa, NJ, pp. 223-243. https://doi.org/10.1007/978-1-62703-389-3_16

Earl DA. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* **4**: 359-361.

Edwards TP, Trigiano R, Wadl P, Ownley BH, Windham A, Hadziabdic D. 2017. First report of *Alternaria alternata* causing leaf spot on whorled sunflower (*Helianthus verticillatus*) in the southeast United States. *Plant Disease* **101**: 632.

Ellis J, Pashley C, McCauley D. 2006. High genetic diversity in a rare and endangered sunflower as compared to a common congener. *Molecular Ecology* **15**: 2345-2355.

Ellis JR. 2008. Conservation genetics of the endangered sunflower *Helianthus verticillatus.* Thesis. Vanderbilt University.

Ellis JR, Bentley KE, McCauley DE. 2008. Detection of rare paternal chloroplast inheritance in controlled crosses of the endangered sunflower *Helianthus verticillatus*. *Heredity* **100**: 574-580.

Ellis JR, McCauley DE. 2009. Phenotypic differentiation in fitness related traits between populations of an extremely rare sunflower: conservation management of isolated populations. *Biological Conservation* **142**: 1836-1843.

Ellstrand NC, Elam DR. 1993. Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology and Systematics* **24**: 217-242.

Ellstrand NC, Roose ML. 1987. Patterns of Genotypic Diversity in Clonal Plant Species. *American Journal of Botany* **74**: 123-131.

ESA. 1973. Endangered Species Act. https://www.fws.gov/endangered/esa-library/pdf/ESAall.pdf

Escaravage N, Questiau S, Pornon A, Doche B, Taberlet P. 1998. Clonal diversity in a *Rhododendron ferrugineum* L. (Ericaceae) population inferred from AFLP markers. *Molecular Ecology* **7**: 975-982.

Esselman E, Jianqiang L, Crawford D, Windus J, Wolfe A. 1999. Clonal diversity in the rare *Calamagrostis porteri* ssp. *insperata* (Poaceae): comparative results for allozymes and random amplified polymorphic DNA (RAPD) and intersimple sequence repeat (ISSR) markers. *Molecular Ecology* **8**: 443-451.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**: 2611-2620.

Excoffier L, Lischer HE. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* **10**: 564-567.

Fernandes PM, Botelho HS. 2003. A review of prescribed burning effectiveness in fire hazard reduction. *International Journal of Wildland Fire* **12**: 117-128.

Flint-Garcia SA, Thornsberry JM, Buckler IV ES. 2003. Structure of linkage disequilibrium in plants. *Annual Review of Plant Biology* **54**: 357-374.

Francis R. 2017. POPHELPER: an R package and web app to analyse and visualize population structure. *Molecular Ecology Resources* **17**: 27-32. doi: 10.1111/1755-0998.12509

Gevaert SD, Mandel JR, Burke JM, Donovan LA. 2013. High genetic diversity and low population structure in Porter's sunflower (*Helianthus porteri*). *Journal of Heredity* **104**: 407-415.

Gitzendanner MA, Soltis PS. 2000. Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany* **87**: 783-792.

Goudet J. 2005. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology Resources* **5**: 184-186.

Groeschl D, Johnson JE, Smith DW. 1992. Early vegetative response to wildfire in a table mountain-pitch pine forest. *International Journal of Wildland Fire* **2**: 177-184.

Grubbs KC, Wynes A. 2015. Reproductive biology of the endangered Schweinitz's sunflower (*Helianthus schweinitzii*). *Castanea* **80**: 20-28.

Hamrick JL, Godt MW. 1990. Allozyme diversity in plant species. *Plant Population Genetics, Breeding, and Genetic Resources*. 43-63. Brown AHD, Clegg MT, Kahler AL, Weir BS, eds. Sinaur, Sunderland.

Heiser CB, Smith DM, Clevenger SB, Martin WC. 1969. The north American sunflowers (*Helianthus*). *Memoirs of the Torrey Botanical Club* **22**: 1-218.

Hill MO. 1973. Diversity and evenness: A unifying notation and its consequences. *Ecology* **54**: 427-432.

Hochkirch A, Adorf F. 2007. Effects of prescribed burning and wildfires on Orthoptera in Central European peat bogs. *Environmental Conservation* **34**: 225-235.

Hubisz MJ, Falush D, Stephens M, Pritchard JK. 2009. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* **9**: 1322-1332.

Jombart T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**: 1403-1405.

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* **11**: 94-109.

Judson OP, Normark BB. 1996. Ancient asexual scandals. *Trends in Ecology & Evolution* **11**: 41-46.

Kamvar Z, Larsen M, Kanaskie A, Hansen E, Grünwald N. 2015. Spatial and temporal analysis of populations of the sudden oak death pathogen in Oregon forests. *Phytopathology* **105**: 982-989.

Kamvar ZN, Amaradasa BS, Jhala R, McCoy S, Steadman JR, Everhart SE. 2017. Population structure and phenotypic variation of *Sclerotinia sclerotiorum* from dry bean (*Phaseolus vulgaris*) in the United States. *PeerJ* **5**: e4152.

Kamvar ZN, Tabima JF, Grünwald NJ. 2014. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**: e281.

Keller LF, Jeffery KJ, Arcese P, Beaumont MA, Hochachka WM, Smith JN, Bruford MW. 2001. Immigration and the ephemerality of a natural population bottleneck: evidence from molecular markers. *Proceedings of the Royal Society of London B: Biological Sciences* **268**: 1387-1394.

Kim D-H, Heber D, Still D. 2004. Genetic diversity of Echinacea species based upon amplified fragment length polymorphism markers. *Genome* **47**: 102-111.

Lienert J. 2004. Habitat fragmentation effects on fitness of plant populations–a review. *Journal for Nature Conservation* **12**: 53-72.

Loveless MD, Hamrick JL. 1984. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* **15**: 65-95.

Luikart G, Sherwin W, Steele B, Allendorf F. 1998. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. *Molecular Ecology* **7**: 963-974.

Mandel JR. 2010. Clonal diversity, spatial dynamics, and small genetic population size in the rare sunflower, *Helianthus verticillatus*. *Conservation Genetics* **11**: 2055-2059.

Mandel JR, Milton EF, Donovan LA, Knapp SJ, Burke JM. 2013. Genetic diversity and population structure in the rare Algodones sunflower (*Helianthus niveus* ssp. *tephrodes*). *Conservation Genetics* **14**: 31-40.

Markell SG, Harveson RM, Block CC, Gulya TJ. 2015. Sunflower diseases. In *Sunflower: Chemistry*, *Production, Processing and Utilization.* 93-128. Force EM, Dunford NT, and Salas JL, eds. Elsevier. American Oilseed Chemists' Society.

Maruyama T, Fuerst PA. 1985. Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics* **111**: 675-689.

Matthews JF, Allison JR, Ware RT, Sr., Nordman C. 2002. *Helianthus verticillatus* Small (Asteraceae) Rediscovered and Redescribed. *Castanea* **67**: 13-24.

Nei M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583-590.

Nei M, Roychoudhury AK. 1974. Sampling variances of heterozygosity and genetic distance. *Genetics* **76**: 379-390. O'Donnell K, Sharrock S. 2017. The contribution of botanic gardens to ex situ conservation through seed banking. *Plant Diversity* **39**: 373-378.

Pashley CH, Ellis JR, McCauley DE, Burke JM. 2006. EST Databases as a Source for Molecular Markers: Lessons from *Helianthus*. *Journal of Heredity* **97**: 381-388.

Piry S, Luikart G, Cornuet J. 1999. BOTTLENECK: a computer program for detecting recent reductions in the effective population size using allele frequency data. *Journal of Heredity* **90**: 502-503.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945-959.

RCoreTeam. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2015. https://www.r-project.org/

Rieseberg LH. 1991. Homoploid reticulate evolution in *Helianthus* (Asteraceae): Evidence from ribosomal genes. *American Journal of Botany*: 1218-1237.

RStudio. 2012. RStudio: integrated development environment for R. *RStudio Inc, Boston, Massachusetts*. https://www.rstudio.com/

Sauve RJ, Mmbaga MT, Zhou S. 2004. In vitro regeneration of the Tennessee coneflower (*Echinacea tennesseensis*). *In Vitro Cellular & Developmental Biology-Plant* **40**: 325-328.

Schmidt LS, Schmidt JO, Rao H, Wang W, Xu L. 1995. Feeding Preference and Survival of Young Worker Honey Bees (Hymenoptera*: Apidae*) Fed Rape, Sesame, and Sunflower Pollen. *Journal of Economic Entomology* **88**: 1591–1595.

Schmitt J. 1980. Pollinator foraging behavior and gene dispersal in Senecio (Compositae). *Evolution* **34**: 934-943.

Schneller J, Holderegger R, Gugerli F, Eichenberger K, Lutz E. 1998. Patterns of genetic variation detected by RAPDs suggest a single origin with subsequent mutations and long-distance dispersal in the apomictic fern *Dryopteris remota* (Dryopteridaceae). *American Journal of Botany* **85**: 1038-1042.

Seiler GJ, Gulya TJ. 2004. Exploration for wild *Helianthus* species in North America: Challenges and opportunities in the search for global treasures. In *International Sunflower Conference Proceedings*, Vol 1, pp. 43-68.

Shannon CE. 2001. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review* **5**: 3-55.

Slatkin M. 2008. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. *Nature Reviews Genetics* **9**: 477-485.

Storfer A. 1999. Gene flow and endangered species translocations: a topic revisited. *Biological Conservation* **87**: 173-180.

Szpiech ZA, Rosenberg NA. 2011. On the size distribution of private microsatellite alleles. *Theoretical Population Biology* **80**: 100-113.

Tian Y, Sun J, Li H, Wang G, Ma Y, Liu D, Quan J, Shan W. 2015. Dominance of a single clonal lineage in the *Phytophthora infestans* population from northern Shaanxi, China revealed by genetic and phenotypic diversity analysis. *Plant Pathology* **64**: 200-206.

Trigiano R, Bernard EC, Hadziabdic D, Dattilo AJ, Wadl PA. 2015. First report of powdery mildew on whorled sunflower (*Helianthus verticillatus*) caused by *Golovinomyces ambrosiae*. *Plant Disease* **200**: 1017. https://doi.org/10.1094/PDIS-11-15-1269-PDN

USFWS. 1991. Endangered and threatened wildlife and plants; *Helianthus schweinitzii* (Schweinitz's sunflower) determined to be endangered. Final Rule May 7, 1991. in: Federal Register Vol. 56 No. 88. In *21088* (ed. I Fish and Wildlife Service), Federal Register.

USFWS. 2002.Endangered and threatened wildlife and plants; removal of *Potentilla robbinsiana* from the federal list of endangered and threatened plants. Final Rule August 27, 2002 in: Federal Register Vol. 67 No. 166. In *54968* (ed. I Fish and Wildlife Service), Federal Register.

USFWS. 2011. Endangered and threatened wildlife and plants; removal of *Echinacea tennesseensis* (Tennessee Purple Coneflower) from the federal list of endangered and threatened plants. Final Rule August 3, 2011 in: Federal Register Vol. 76 No. 149. In *46632* (ed. I Fish and Wildlife Service), Federal Register.

USFWS. 2014. Endangered and threatened wildlife and plants; designation of critical Habitat for *Physaria globosa* (Short's bladderpod), *Helianthus verticillatus* (whorled sunflower), and *Leavenworthia crassa* (fleshy-fruit gladecress). Final Rule August 26, 2014. in: Federal Register Vol. 79 No. 165.. In *50990* (ed. I Fish and Wildlife Service), Federal Register.

Volis S. 2017. Conservation utility of botanic garden living collections: Setting a strategy and appropriate methodology. *Plant Diversity* **39**: 365-372.

Walck JL, Hemmerly TE, Hidayati SN. 2002. The endangered Tennessee purple coneflower *Echinacea tennesseensis* (Asteraceae): its ecology and conservation. *Native Plants Journal* **3**: 54-64.

Wang X, Rinehart TA, Wadl PA, Spiers JM, Hadziabdic D, Windham MT, Trigiano RN. 2009. A new electrophoresis technique to separate microsatellite alleles. *African Journal of Biotechnology* **8**: 2432-2436.

Whelan RJ. 2002. Managing fire regimes for conservation and property protection: an Australian response. *Conservation Biology* **16**: 1659-1661.

Widén B, Andersson S. 1993. Quantitative genetics of life-history and morphology in a rare plant, *Senecio integrifolius*. *Heredity* **70**: 503.

Willi Y, Buskirk JV, Hoffmann AA. 2006. Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics* **37**: 433-458.

Wyatt R. 1992. Conservation of rare and endangered bryophytes: Input from population genetics. *Biological Conservation* **59**: 99-107.

Young A, Boyle T, Brown T. 1996. The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution* **11**: 413-418.

# Appendix

**Table 1.** Fourteen microsatellite loci used to examine genetic diversity and spatial population structure of *Helianthus verticillatus* in Georgia and Tennessee, USA.

| Core Loci[A] | Renamed Loci[B] | Forward and reverse primers 5'-3' | Repeat motif | Number of alleles | Allele size range (bp) |
|---|---|---|---|---|---|
| BL0006 | HV006 | F: CATGGGTGATCAATGGAGTG<br>R: CGGCACATAACAAGTGCTTC | (gtga)3 | 14 | 225-279 |
| BL0013 | HV012 | F: CGAGACGGTTAAGAGCTTGC<br>R: GGTGTACAACCAACTCACACC | (gtta)3 | 16 | 319-364 |
| BL0022 | HV017 | F: ACTTACCGTTGCATTTGGTG<br>R: GCTTATCCCTAGAACACGATTACAG | (taa)4 | 3 | 105-111 |
| BL0015 | HV024 | F: AATTGGAGCGGATGGTATTG<br>R: AATATCTCTTATTTCAATAGTCCAACG | (atg)4 | 5 | 356-369 |
| BL0019 | HV026 | F: GAGTCCTGGCCTGAACAGAG<br>R: CAAACTGCAATGTACCTTCTTGAC | (gaaa)3 | 8 | 292-316 |
| BL0024 | HV028 | F: CTCCCGCACTTCAAGCTAAC<br>R: CATACACCTTTGCGGTTTCC | (gtaa)3 | 5 | 117-126 |
| BL0031 | HV031 | F: CCGGAAGATAACGACGAGTG<br>R: TCCATCGCTTTCCCTAAATC | (gac)4 | 10 | 405-437 |
| BL0033 | HV033 | F: GGGAGTTACACGCCTCCAG<br>R: CACAACCATACGCCATCAAG | (cac)4 | 5 | 270-284 |
| BL0034 | HV034 | F: GGTCGTCTACTACGGCTTCG<br>R: TAACCGAACGACCATTCTTC | (tgtt)4 | 4 | 155-165 |
| BL0037 | HV037 | F: GGTTAGGGTGAGGGTGGTG<br>R: AAGCCATAGTAAGTTCCTCTTACAAAC | (tgca)3 | 7 | 153-179 |
| BL0041 | HV041 | F: ACATTTGGACGTTTGGAAGC<br>R: TCCATCGAGATGTTGACACG | (ctt)4 | 2 | 185-189 |
| BL0042 | HV042 | F: GGTTACAACGGTGGAAGTCG<br>R: TCCGGTTCACCAATTCATTC | (ggc)4 | 16 | 364-430 |
| BL0046 | HV046 | F: GAACCAACACAACCAAATCC<br>R: TGTCGCTTCAACGCATAAAC | (aaca)3 | 10 | 312-339 |
| BL0048 | HV048 | F: TTGTGGAGACGGTGAATGAG<br>R: TAACCGAACGACCATTCTTC | (gaa)4 | 5 | 215-233 |

[A] – All core loci taken from Pashley et al. (2006); [B] – Loci renamed for the purpose of this study only

**Table 2.** Genetic diversity indices of *Helianthus verticillatus* samples from two sampling sites (C and R populations) analyzed as fourteen collection zones using fourteen microsatellite loci.

| Collection Zone | N | MLG | H | Pa | $H_{exp}$ | $\bar{r}_d$ | *P-value* ($\bar{r}_d$) |
|---|---|---|---|---|---|---|---|
| C1 | 27 | 26 | 3.26 | 0 | 0.44 | 0.15 | *P*<0.001 |
| C2 | 17 | 17 | 2.83 | 2 | 0.41 | 0.15 | *P*<0.001 |
| C3 | 5 | 5 | 1.61 | 0 | 0.54 | 0.25 | *P*<0.001 |
| C4 | 11 | 11 | 2.4 | 0 | 0.43 | 0.11 | *P*<0.001 |
| C5 | 14 | 13 | 2.56 | 0 | 0.36 | 0.23 | *P*<0.001 |
| R1 | 19 | 19 | 2.94 | 3 | 0.54 | 0.08 | *P*<0.001 |
| R2 | 19 | 19 | 2.94 | 1 | 0.51 | 0.05 | *P*<0.001 |
| R3 | 17 | 17 | 2.83 | 2 | 0.55 | 0.09 | *P*<0.001 |
| R4 | 18 | 18 | 2.89 | 1 | 0.51 | 0.1 | *P*<0.001 |
| R5 | 19 | 19 | 2.94 | 3 | 0.52 | 0.08 | *P*<0.001 |
| R6 | 14 | 14 | 2.64 | 0 | 0.45 | 0.03 | *P*=0.07 |
| R7 | 8 | 7 | 1.95 | 0 | 0.57 | 0.28 | *P*<0.001 |
| R8 | 8 | 8 | 2.08 | 2 | 0.59 | 0.13 | *P*<0.001 |
| R9 | 10 | 10 | 2.3 | 7 | 0.54 | 0.09 | *P*<0.001 |
| **Total/Mean** | **206** | **203** | **2.58** | **21** | **0.50** | **0.13** | |

N - total number of samples; MLG - number of multi locus genotypes observed after clone correction; H - Shannon-Wiener index of MLG diversity; Pa - number of private alleles in each population; $H_{exp}$ - Nei's genotypic diversity corrected for sample size; $\bar{r}_d$ - the standardized index of association.

**Table 3**. Analysis of molecular variance (AMOVA) for *Helianthus verticillatus* across fourteen microsatellite loci for the all collection zones structured as one hierarchal group (A), two groups separated by sampling site (B), two clusters as indicated by STRUCTURE (C), and two clusters as indicated by discriminate analysis of principle components (DAPC) results (D).

| Source of Variation | d.f. | Sum of Squares | Variance Components | Percentage of Variation | *P* Value |
|---|---|---|---|---|---|
| **A.** Analysis of 14 Collection zones as one hierarchal group | | | | | |
| **Among Collection Zones** | 13 | 449.39 | 1.08 Va | 23.96 | *P*<0.001 |
| **Within Collection Zones** | 394 | 1351.7 | 3.43 Vb | 76.04 | *P*<0.001 |
| **Total** | 407 | 1801.1 | 4.51 | | |
| Fixation indices: Fst=0.24 | | | | | |
| **B.** Analysis of two sampling sites (C and R sites) | | | | | |
| **Among Collection Sites** | 1 | 181.82 | 0.93 Va | 18.87 | *P*<0.001 |
| **Among Individuals within Collection Sites** | 202 | 1107.3 | 1.49 Vb | 30.17 | *P*<0.001 |
| **Within Individuals** | 204 | 512 | 2.50 Vc | 50.96 | *P*<0.001 |
| **Total** | 407 | 1801.1 | 4.93 | | |
| Fixation indices: Fst=0.19, Fis=0.37, Fit=0.49 | | | | | |
| **C.** Analysis of 2 clusters as indicated by STRUCTURE results | | | | | |
| **Among Clusters** | 1 | 206.02 | 0.99 Va | 20.18 | *P*<0.001 |
| **Among Individuals within Clusters** | 202 | 1083.1 | 1.43 Vb | 28.92 | *P*<0.001 |
| **Within Individuals** | 204 | 512 | 2.51 Vc | 50.99 | *P*<0.001 |
| **Total** | 407 | 1801.1 | 4.93 | | |
| Fixation indices: Fst=0.20, Fis=0.36, Fit=0.49 | | | | | |

Fst - variance among sampling sites relative to the total variance; Fis - inbreeding coefficient of individuals relative to the population; Fit - variance in the total population

**Table 3 cont.**

| Source of Variation | d.f. | Sum of Squares | Variance Components | Percentage of Variation | *P* Value |
|---|---|---|---|---|---|
| **D.** Analysis of 2 clusters as indicated by DAPC results | | | | | |
| **Among Clusters** | 1 | 210.04 | 1.02 Va | 20.67 | *P*<0.001 |
| **Among Individuals within Clusters** | 202 | 1079.05 | 1.42 Vb | 28.61 | *P*<0.001 |
| **Within Individuals** | 204 | 512 | 2.51 Vc | 50.72 | *P*<0.001 |
| **Total** | 407 | 1801.09 | 4.95 | | |
| Fixation indices: Fst=0.21, Fis=0.36, Fit=0.49 | | | | | |

Fst - variance among sampling sites relative to the total variance; Fis - inbreeding coefficient of individuals relative to the population; Fit - variance in the total population

**Table 4.** Pairwise population matrix of Nei's unbiased genetic distance of *Helianthus verticillatus* samples from two sampling sites (C and R) split into fourteen collection zones.

| | C1 | C2 | C3 | C4 | C5 | R1 | R2 | R3 | R4 | R5 | R6 | R7 | R8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **C2** | 0.14 | – | | | | | | | | | | | |
| **C3** | 0.25 | 0.12 | – | | | | | | | | | | |
| **C4** | 0.33 | 0.18 | 0.15 | – | | | | | | | | | |
| **C5** | 0.32 | 0.18 | 0.17 | 0.02 | – | | | | | | | | |
| **R1** | 0.26 | 0.22 | 0.17 | 0.19 | 0.18 | – | | | | | | | |
| **R2** | 0.66 | 0.66 | 0.60 | 0.50 | 0.57 | 0.33 | – | | | | | | |
| **R3** | 0.56 | 0.56 | 0.53 | 0.43 | 0.48 | 0.35 | 0.14 | – | | | | | |
| **R4** | 0.66 | 0.59 | 0.53 | 0.45 | 0.51 | 0.38 | 0.14 | 0.10 | – | | | | |
| **R5** | 0.60 | 0.56 | 0.51 | 0.39 | 0.44 | 0.30 | 0.09 | 0.09 | 0.06 | – | | | |
| **R6** | 0.64 | 0.62 | 0.58 | 0.45 | 0.51 | 0.35 | 0.08 | 0.14 | 0.11 | 0.08 | – | | |
| **R7** | 0.52 | 0.53 | 0.61 | 0.62 | 0.68 | 0.44 | 0.18 | 0.22 | 0.23 | 0.23 | 0.17 | – | |
| **R8** | 0.89 | 1.06 | 0.84 | 0.92 | 0.99 | 0.68 | 0.42 | 0.38 | 0.46 | 0.44 | 0.42 | 0.37 | – |
| **R9** | 0.93 | 0.92 | 0.86 | 0.95 | 1.03 | 0.91 | 0.73 | 0.70 | 0.70 | 0.77 | 0.69 | 0.52 | 0.27 |

**Table 5.** Pairwise population differentiation for *Helianthus verticillatus* samples from two sampling sites (C and R) divided into fourteen subpopulations using 14 microsatellite loci. Fst values are based on 9,999 permutations.

| | C1 | C2 | C3 | C4 | C5 | R1 | R2 | R3 | R4 | R5 | R6 | R7 | R8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **C2** | 0.09 | – | | | | | | | | | | | |
| **C3** | 0.08 | 0.05 | – | | | | | | | | | | |
| **C4** | 0.12 | 0.06 | 0.05 | – | | | | | | | | | |
| **C5** | 0.16 | 0.09 | 0.08 | 0.02 | – | | | | | | | | |
| **R1** | 0.11 | 0.10 | 0.05 | 0.07 | 0.09 | – | | | | | | | |
| **R2** | 0.21 | 0.21 | 0.12 | 0.18 | 0.22 | 0.12 | – | | | | | | |
| **R3** | 0.17 | 0.18 | 0.12 | 0.17 | 0.21 | 0.12 | 0.07 | – | | | | | |
| **R4** | 0.21 | 0.20 | 0.12 | 0.17 | 0.21 | 0.13 | 0.07 | 0.06 | – | | | | |
| **R5** | 0.19 | 0.19 | 0.11 | 0.15 | 0.19 | 0.11 | 0.06 | 0.05 | 0.04 | – | | | |
| **R6** | 0.22 | 0.25 | 0.17 | 0.23 | 0.27 | 0.15 | 0.08 | 0.08 | 0.08 | 0.06 | – | | |
| **R7** | 0.15 | 0.19 | 0.18 | 0.23 | 0.26 | 0.13 | 0.09 | 0.08 | 0.09 | 0.08 | 0.06 | – | |
| **R8** | 0.19 | 0.25 | 0.19 | 0.25 | 0.28 | 0.16 | 0.13 | 0.10 | 0.13 | 0.11 | 0.13 | 0.10 | – |
| **R9** | 0.23 | 0.28 | 0.21 | 0.28 | 0.31 | 0.21 | 0.20 | 0.18 | 0.19 | 0.19 | 0.20 | 0.15 | 0.09 |

**Table 6.** Bottleneck determination by sign tests for *Helianthus verticillatus* samples using fourteen microsatellite loci and grouped by STRUCTURE results (A), by sampling site (B) and by discriminate analysis of principle components (DAPC) results (C).

**A.** Bottleneck results when grouped by STRUCTURE results.

| | Mutation Model (Excess/Deficit)[A] | | | | |
|---|---|---|---|---|---|
| **STRUCTURE Group** | **IAM** | **TPM** | **SMM** | **Mode-Shift[B]** | ***P*-value** |
| 1 | **14**/0 | **14**/0 | **12**/2 | Shifted | *P*<0.01 |
| 2 | **14**/0 | **14**/0 | **12**/2 | Shifted | *P*<0.01 |

**B.** Bottleneck results when grouped by collection zone.

| | Mutation Model (Excess/Deficit)[A] | | | | |
|---|---|---|---|---|---|
| **Collection Zone** | **IAM** | **TPM** | **SMM** | **Mode-Shift[B]** | ***P*-value** |
| R | **14**/0 | **14**/0 | **12**/2 | Shifted | *P*<0.01 |
| C | **14**/0 | **14**/0 | **12**/2 | Shifted | *P*<0.01 |

**C.** Bottleneck results when grouped by DAPC results.

| | Mutation Model (Excess/Deficit)[A] | | | | |
|---|---|---|---|---|---|
| **DAPC Group** | **IAM** | **TPM** | **SMM** | **Mode-Shift[B]** | ***P*-value** |
| 1 | **14**/0 | **14**/0 | **12**/2 | Shifted | *P*<0.01 |
| 2 | **14**/0 | **14**/0 | **12**/2 | Shifted | *P*<0.01 |

I.A.M. = infinite allele model; T.P.M. = two-phase mutation model; S.M.M. = stepwise mutation model. [A] – Excess/deficit indicates the number of loci showing excess/deficit of gene diversity under mutation-drift equilibrium; [B] – A shift in the distribution of allelic frequency classes is expected in populations that experienced recent bottleneck event.

**Table 7.** Comparing sample size, genetic diversity and genetic structure from multiple species of *Helianthus* in past studies with data generated in this study on *H. verticillatus*.

| Species | Sample Size (N) | Number of Loci Tested | Genetic Diversity ($H_{exp}$) | Genetic Structure ($F_{st}$) |
|---|---|---|---|---|
| *H. angustifolius*[A] | 48 | 11 | 0.34 | 0.17 |
| *H. annus*[A] | 12 | 11 | 0.58 | N/A |
| *H. grosseseratus*[A] | 56 | 11 | 0.44 | N/A |
| *H. nieveus* spp. *tephrodes*[A] | 119 | 11 | 0.31 | 0.17 |
| *H. porteri*[A] | 200 | 18 | 0.62 | 0.12 |
| *H. verticillatus*[A] | 71 | 11 | 0.48 | 0.12 |
| *H. verticillatus*[B] | 206 (203[C]) | 14 | 0.50 | 0.24 |

[A] – Data taken from Mandel et al. (2013) and Gevaert et al (2013); [B] – Data generated in this study; [C] – Number of samples after clone correction

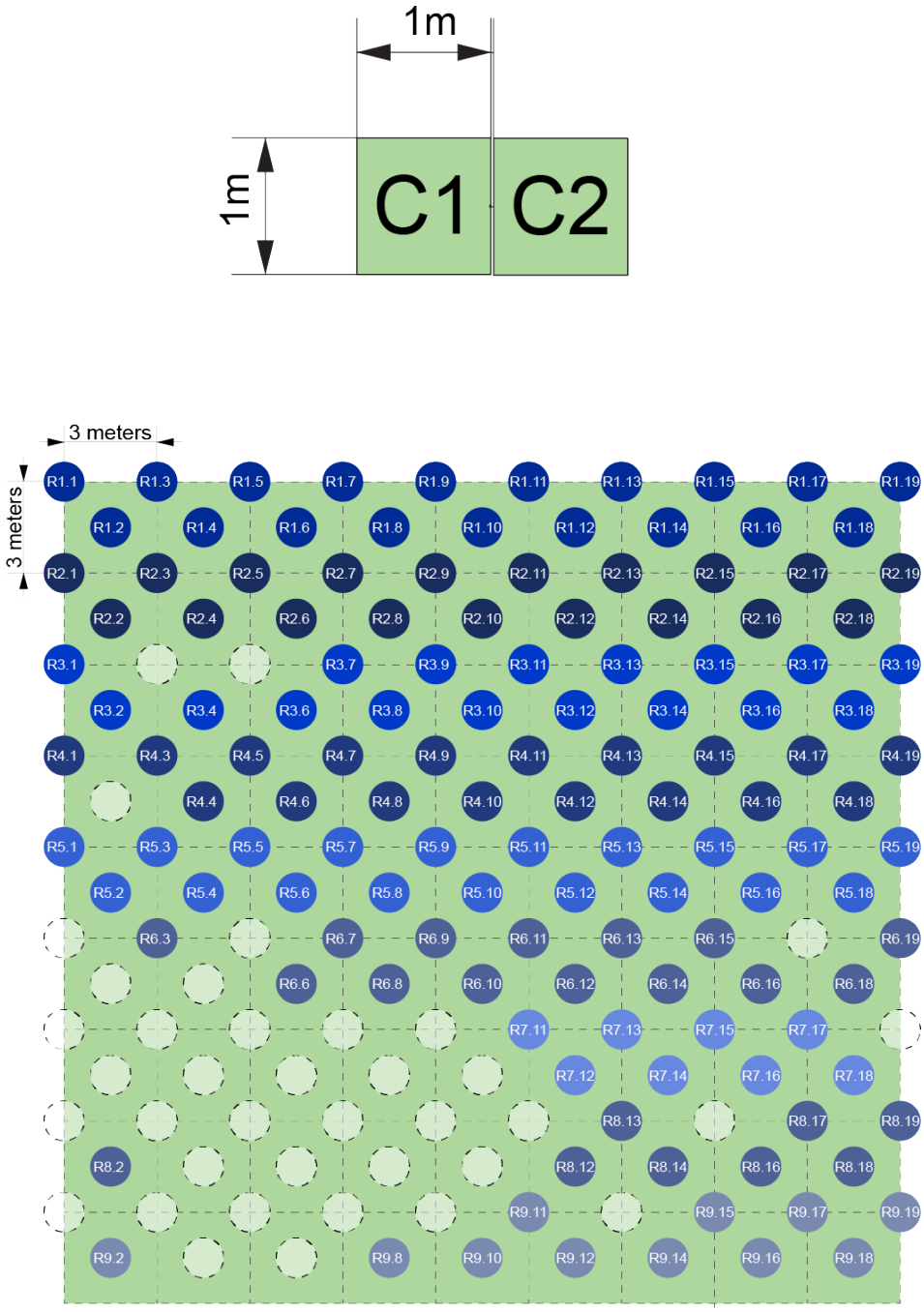


**Figure 1** Diagram showing sampling (grid) methods employed at sites one (upper) and two (lower). Site one consisted of five contiguous 1 meter x 1 meter plots (only two of which are represented in the diagram). At site two, a 3 meters x 3 meters grid was created. The circles represent sampling locations within the grid. Sample names (RX.X) are inlayed within each circle and collection zones (R1-9) are differentiated by color. Empty circles represent locales where no plant was present and no sample was taken (lower diagram).

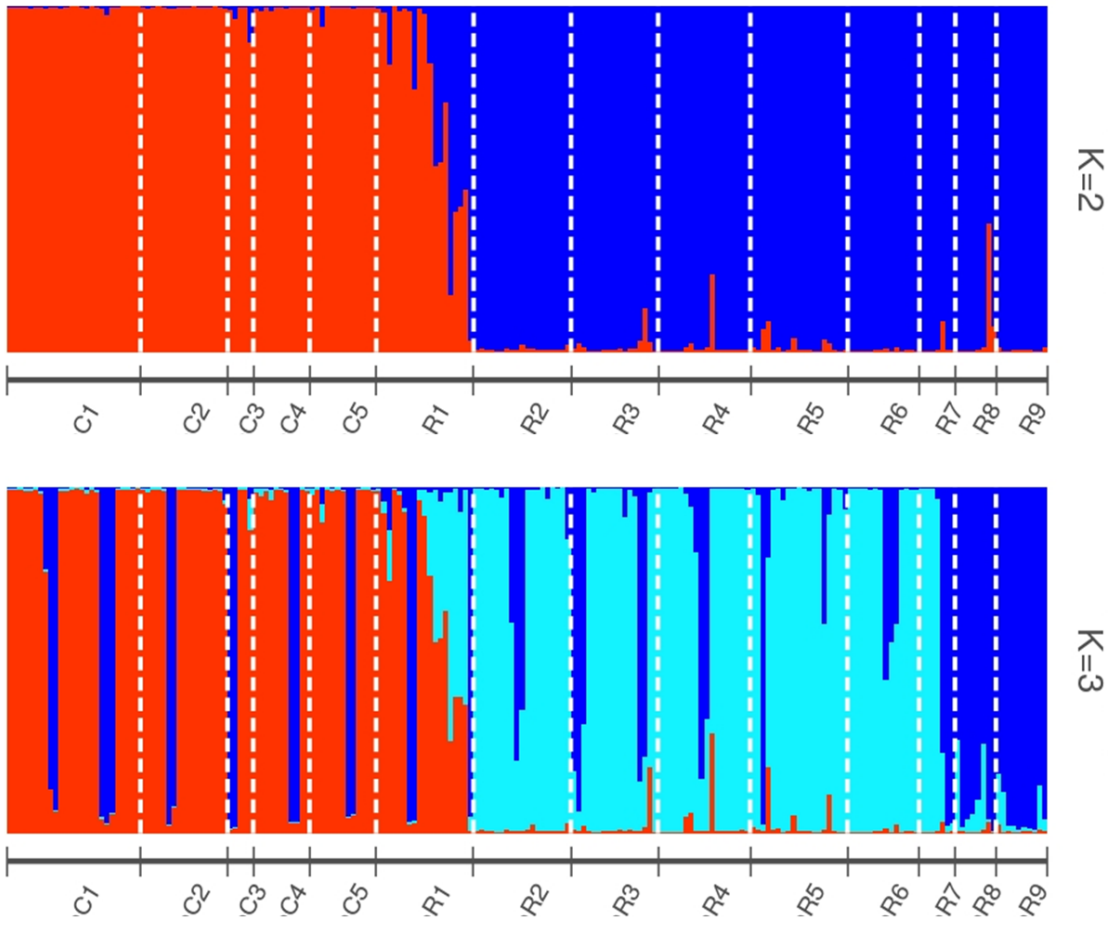


**Figure 2.** STRUCTURE bar graphs representing genetic clusters (K=2-3) of samples from the two sampling sites (C and R) divided into fourteen collection zones. Each bar represents an individual sample and colors code membership of each sample's assigned cluster. Using Evanno's method, the presence of two genetic clusters (K=2) was found to be the best fit for this data.

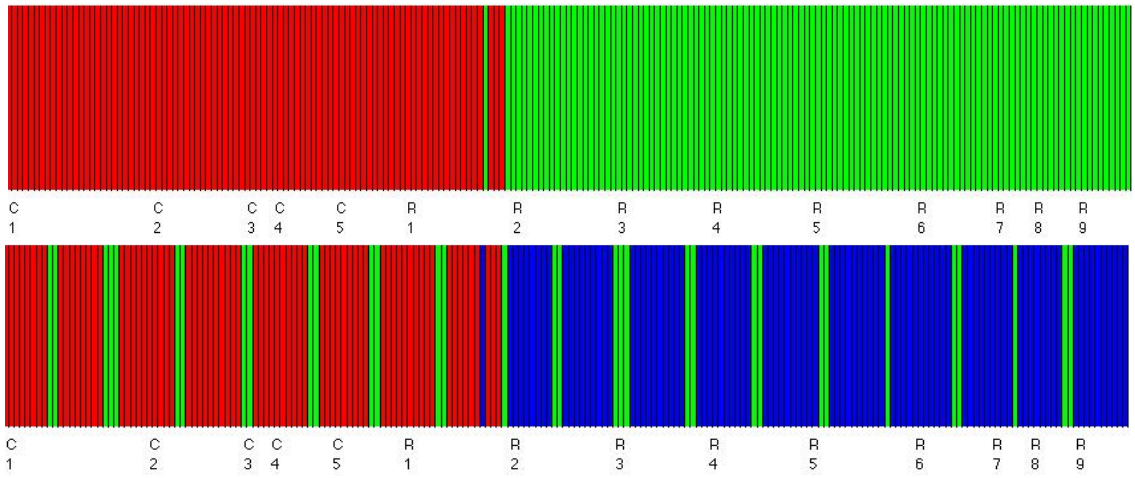


**Figure 3.** Population structure and clustering for *Helianthus verticillatus* by collection zone using the Bayesian analysis program BAPS. The first graph (A) is assuming two clusters and the second graph (B) is assuming three clusters.

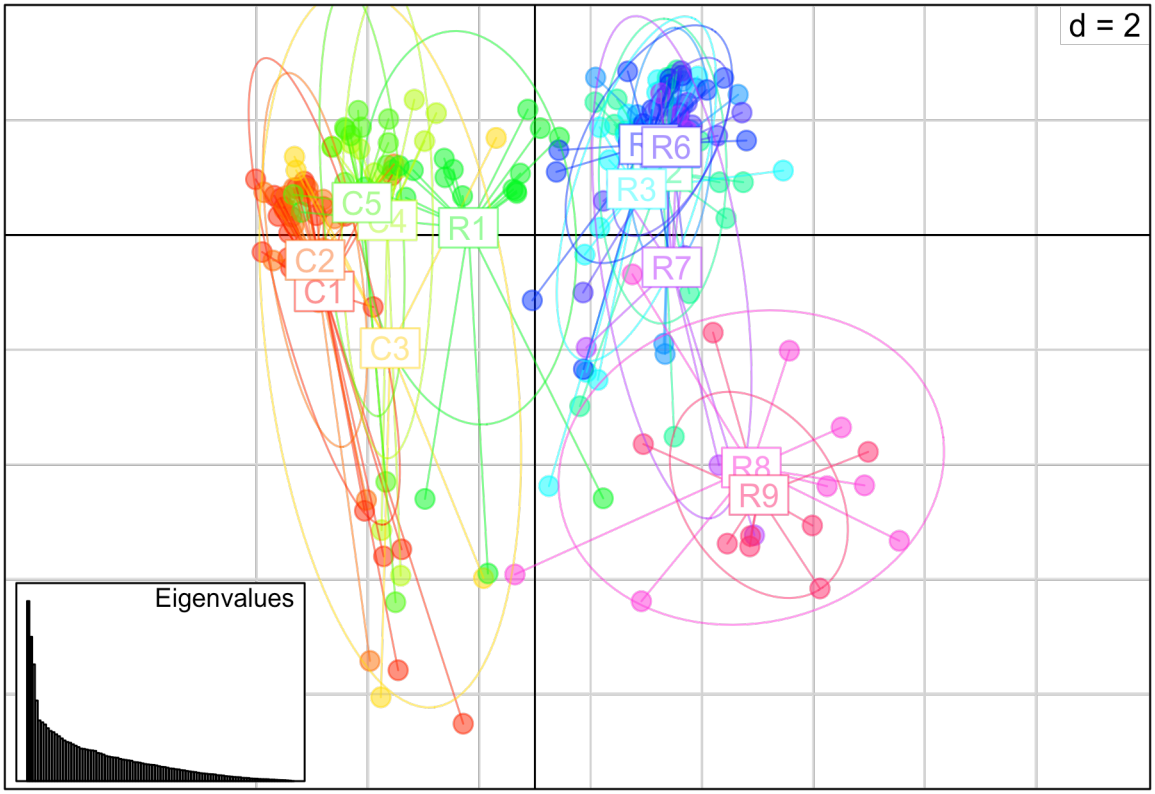


**Figure 4.** Principle coordinates analysis (PCoA) using discriminant analysis of principle components (DAPC) method among fourteen collection zones of *Helianthus verticillatus* samples using fourteen microsatellite loci.

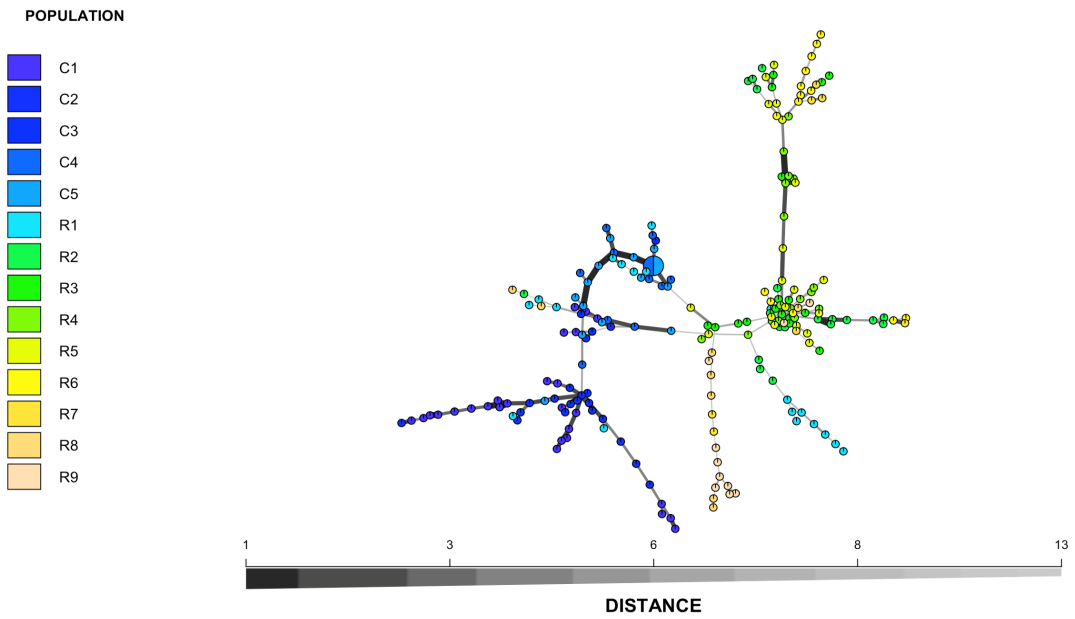


**Figure 5.** Minimum spanning network (MSN) of *Helianthus verticillatus* based on Bruvo's genetic distance for fourteen microsatellite loci. The nodes of the MSN represent individual multilocus genotypes (MLG) with the size and color representing population membership size and designated collection zone, respectively. To avoid overlapping nodes, the size was scaled to log1.75n, where n equals the node sample size. Lines between nodes represent genetic distance between MLG as determined by Prim's algorithm.

**Vita**

Tyler Edwards was born in Kingsport, TN and grew up in Maces Springs, VA. He graduated cum laude from Virginia Highlands Community College with an A.S. in Science focusing in Horticulture in 2012, then cum laude from Virginia Tech with a B.S. in Environmental Horticulture in 2014. After finishing his B.S. degree, he worked in the Department of Plant Pathology, Physiology and Weed Science at Virginia Tech the following year focusing primarily on boxwood blight as well as downy and powdery mildew in wine grapes. He started his Masters in Entomology and Plant Pathology in the Hadziabdic lab at the University of Tennessee in 2015. Here he worked on conservation of the rare species *Helianthus verticillatus* and *Pityopsis ruthii*.